DOCUMENT RESUME

ED 420 521 SE 061 560

AUTHOR Tien, L. T.; Stacy, A. M.
TITLE Promoting Scientific Inquiry through the MORE Laboratory.
PUB DATE 1998-04-00
NOTE 22p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Diego, CA, April 13-17, 1998).
PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE EDRS Price MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS *Chemistry; Classroom Environment; Cognitive Processes; Computer Uses in Education; Concept Formation; Educational Experience; Higher Education; *Inquiry; *Laboratory Experiments; *Problem Solving; Science Activities; Science Curriculum; Science Education; Undergraduate Study

ABSTRACT

The purpose of this study is to investigate the effects of two laboratory environments for introductory chemistry on undergraduate students' inquiry skills. The two laboratory environments are embedded in the same lecture course. The standard laboratory curriculum and instruction are distinct from the curriculum explored in this study, the MORE (Model, Observe, Reflect, Explain) curriculum. The MORE curriculum emphasizes how to think through the inquiry process rather than focusing on how to perform laboratory procedures and algorithmic calculations. This study examines the classroom interactions in each setting and analyzes students' abilities to design their own experiments. Analysis of videotapes shows a marked contrast between the two laboratory environments. Interaction in the MORE environment focuses on a reflective level and the students in that environment appear to make connections between actions and models more easily. Contains 24 references. (DDR)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 420 521

L. Tien

2A

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Promoting Scientific Inquiry through the MORE Laboratory

by
L.T. Tien
A.M. Stacy

SE 061 560

Promoting scientific inquiry through the MORE laboratory

L. T. Tien & A. M. Stacy
University of California, Berkeley

Abstract

The purpose of this study is to investigate the effects of two laboratory environments for introductory chemistry on undergraduate students' inquiry skills. The two laboratory environments are embedded in the same lecture course, however the standard laboratory curriculum and instruction are distinct from the MORE (Model-Observe-Reflect-Explain) curriculum that we developed. The MORE curriculum emphasizes how to think through the inquiry process, rather than focusing on how to perform laboratory procedures and algorithmic calculations. We studied the classroom interactions in each setting and analyzed students' ability to design their own experiments.

Introduction

The undergraduate chemistry community has recently recognized the need for curricular reform. The chemistry classroom has undergone relatively few changes in the past twenty years (except for increasing the number of concepts covered in the first-year survey course). Currently, the National Science Foundation is funding five systemic change projects aimed at sweeping changes in the way that college chemistry is taught. These projects have focused their energies to develop new curricular materials, implement alternative methods in the classroom, develop computer tools to aid visualization of chemical processes, and assess the impact of the innovations.

Educational research has demonstrated that the typical classroom experience falls short of promoting a deep conceptual understanding, expert problem solving abilities, and inquiry skills (e.g. Arons, 1993; Gottfried, Hoots, Creek, Tamppari, Lord, and Sines, 1993; Spencer, 1993-1994). One observation is that typical college science instruction does not promote an accurate view of science as a discipline and a practice. That is, the experiences of students in the classroom is very different from the actions and thought processes of scientists in a research laboratory. One possible approach to bridge the gap between the typical classroom experience and expert practice is to provide students with a classroom experience that more closely resembles the intellectually stimulating environment of a research scientist. In order to promote a model of authentic practice, we focused on the general chemistry laboratory curriculum to give students a more integrated sense of the domain, and opportunities to engage in model-based reasoning and to design their own experiments.

Theoretical perspective

According to the American Chemical Society's Task Force on Chemistry Education, one of the current shortcomings of laboratory instruction is the lack of development of inquiry skills, that is, the expert thinking underlying the experimentation process (Lloyd and Spencer, 1994). Inquiry skills have also been referred to as the scientific method and science process skills. Basic inquiry skills, such as observation, classification, measurement, prediction, and inference, provide a foundation for learning the more complex integrated skills, such as data analysis and interpretation, designing experiments, and formulating models (Funk, Okey, Fiel, Jans, and Sprague, 1979; Yap and Yeany, 1988; Padilla, 1991). Unfortunately, development of such cognitive skills is oftentimes overshadowed by emphasis on procedural competence.

Scientific inquiry is a cognitively complex, dynamic process. Scientists develop models of understanding a system, with the realization that these models are subject to change over

time. These predictive models are then tested as scientists design appropriate methods and acquire data so that aspects of their model can be refined, elaborated, or eliminated (Koslowski, 1996). The processes of designing experiments, interpreting evidence, and drawing inferences require reflection and self-monitoring (Brown & Campione, 1993; Kuhn, 1993; White & Frederiksen, 1995).

The goals of the science laboratory have ranged from learning to act like a scientist (i.e., learning laboratory skills) to learning to think like a scientist (Kreitler & Kreitler, 1974; Tobin, 1990). However, undergraduate laboratory experiences typically lack the emphasis on learning to think like a scientist. Instead, the undergraduate laboratory has traditionally focused on a narrow portion of the inquiry experience, specifically, collecting data and learning laboratory techniques. Students typically follow prescribed procedures step-by-step, and thus employ minimal experimental design and analysis skills in these settings. In a one-week investigation, students either possess preconceived notions of the outcome or are collecting data to plug into a given formula, and thus, need not grapple with the experimental data to make sense of it or draw conclusions. Students are traditionally given experimental protocols telling them *what to do* in the lab. It is far less common for students to be guided in *how to think* during the inquiry process.

In order to respond to the differences between the standard laboratory experience and expert research practice, the MORE Project laboratory curriculum has been developed for the general chemistry course at the University of California at Berkeley. This curriculum provides students with a model of expert research (Model-Observe-Reflect-Explain) to explore conceptually-rich systems through authentic research opportunities, such as the opportunity to design their own experiments and modify explanatory models. Students invest three to four weeks on each of three modules and investigate problems situated in real-world contexts: "The Chemistry of Smell: How does the nose know?", "The Chemistry of the Stomach: How do YOU spell relief?", and "Determining the Effectiveness of Sunscreen: How can you avoid getting burned?".

In developing the laboratory curriculum, we used guidelines for instructional design garnered from educational research. Specifically, we felt it was important to include the following components: 1) exploration of concepts through authentic scientific inquiry; 2) promotion of metacognition; and 3) support for guided discovery.

First, a classroom experience that more closely resembles that of an expert scientist appears to be more effective in teaching inquiry skills (Carey, Evans, Honda, Jay, and Unger, 1989; Schauble, Glaser, Duschl, Schulze, and John, 1995). The MORE curriculum strives to be "authentic" in the following ways: (a) situating the investigations in conceptually-rich domains, specifically, biologically-relevant systems; (b) providing authentic research opportunities, such as the opportunity to design experiments and refine one's understanding of an experimental system; and (c) giving students guidelines for how to think during the inquiry process as they construct their scientific understanding through the use of the MORE (Model-Observe-Reflect-Explain) Cycle, a model of expert research.

The second component guiding the instructional design is the promotion of metacognition. Metacognition is defined as the awareness of and monitoring of one's thinking (Flavell, 1976; Brown, Bransford, Ferrara, and Campione, 1983). Our primary tool for promoting metacognition is through the use of the MORE Cycle (Tien, Rickey, and Stacy, in press). The Model-Observe-Reflect-Explain Cycle offers students an explicit structure to monitor their thinking during their laboratory experience, both in and out of the classroom. The MORE Cycle is a representation of the thought processes of scientists that encourages

students to reflect upon their own understanding of chemistry throughout the laboratory program.

The third aspect, support for guided discovery, is achieved both through the structure of the laboratory modules and through the instructional methods employed by the teaching assistant. By structuring the modules such that there is increasing responsibility placed on the student, we ensure that students have a basic understanding of the chemical concepts and techniques relevant to the module before attempting to design their own experiments. In addition, the instructor is key to providing the support in the classroom to carry out the goals and philosophy of the MORE project.

> *It may be that what is crucial is for teachers to function as models of rational inquirers and for them to exhibit the practices and values of inquiry in their teaching. Effective modeling, no doubt, will involve both observation and discourse, this borders on the self-evident.* (Strike and Posner, 1992).

Using a cognitive apprenticeship model of instruction, the instructor needs to make expert thinking visible to the students in order for them to adopt those thinking processes into their own process of inquiry (Collins, Brown, and Holum, 1991; Collins, Brown, and Newman, 1989; Tien and Stacy, 1998). Past research has shown that innovations in curriculum cannot be successful when the philosophy and goals underlying such innovations are not understood by the teachers (Baird, Fensham, Gunstone, and White, 1991; Welch, 1979).

Based on this framework of instructional design, the MORE curriculum and its implementation were designed to encourage students to develop progressively more sophisticated models based on the evidence they collect throughout a module. As the module progresses, students are refining their understanding about the system. Within each of the three modules, students are responsible for designing and carrying out an experiment. Students are required to incorporate the concepts and techniques they learned in the early weeks of the module to investigate an experimental question related to the overarching question driving the module. Because the outcomes are not known, students take on ownership of their projects. Also, as the semester progresses, students are given increasing responsibility to carry out the experimental designs and analyses. Thus, one of our hopes for the MORE curriculum was to provide students with an authentic inquiry experience and to have those experiences translate into inquiry skills.

Methodology

In order to assess the impact of the MORE curriculum and instruction as compared to a more traditional laboratory curriculum, an experiment was designed such that students in the test and control groups would be as similar as possible. First, the students did not self-select to experience the MORE laboratory curriculum. The MORE Project laboratory curriculum was implemented in two of the forty-six Chemistry 1A sections during the Fall 1996 semester. Students were randomly assigned to one section among those offered during the time period selected by the student. Thus, the students participating in the MORE curriculum did not know in advance that they would be in a special section. However, on the first day of lab these students were informed of the unique nature of their laboratory curriculum and given the option to switch into a class using the standard lab curriculum. Only one student opted to leave.

Second, the potential instructor effect was addressed by selecting the most experienced and enthusiastic graduate students. Both of the MORE laboratory sections were taught by an experienced graduate student instructor who was involved in the development of the MORE

curriculum. Two other experienced graduate student instructors who were enthusiastic about teaching laboratory sections in the standard way were chosen to teach Control sections.

Third, students with the same intended major typically cluster in similar time periods. Each Control group (Sections 3 and 4) was selected randomly from sections meeting at the same time as each of the MORE sections (Sections 1 and 2). Matching sections within a time period theoretically minimized any pre-existing differences between the populations. Finally, students in all laboratory sections attended lectures given by the same professor.

The MORE laboratory curriculum was compared with a more typical laboratory curriculum implemented in Control sections. The Control laboratory curriculum was comprised of seven one-week laboratory experiments and one multiple-week experiment that did not have the MORE cycle embedded. Like the MORE curriculum, the majority of the laboratory experiments for the Control sections employed a real-world context such as water chemistry and the thermochemistry of foods so context was not a variable. (See Table 1.)

<table>
<tr><th>Week</th><th>Control</th><th>MORE</th></tr>
<tr><td>1</td><td>Model Airbags [C, ED]</td><td rowspan="3">Chemistry of Smells [C, ED]</td></tr>
<tr><td>2</td><td>Gas Laws</td></tr>
<tr><td>3</td><td>The Nose and VSEPR [C]</td></tr>
<tr><td>4</td><td>Chromatography</td><td rowspan="3">Chemistry of the Stomach [C, ED]</td></tr>
<tr><td>5</td><td>Thermo in a nutshell [C, ED]</td></tr>
<tr><td>6</td><td>Minerals, Dry Lakes, and Borax [C]</td></tr>
<tr><td>7</td><td rowspan="3">Natural waters [C, ED]</td><td rowspan="3">Chemistry of Sunscreens [C, ED]</td></tr>
<tr><td>8</td></tr>
<tr><td>9</td></tr>
<tr><td>10</td><td>Spectroscopy [C, ED]</td><td></td></tr>
</table>

Table 1. Laboratory format for MORE and Control curriculum
[C] real-world context; [ED] experimental design component

In addition, the Control students were given some limited experimental design opportunities. However, in contrast to the MORE curriculum, these opportunities did not allow the Control students to formulate new experimental questions or to build upon experience and understanding gained from investigations they had carried out in previous weeks. Finally, instead of emphasizing an evolving conceptual understanding of a system, the Control pre- and post-laboratory assignments focused primarily on procedures, calculations, and pre-defined data analyses.

Students in the MORE and Control classes were administered surveys at the beginning and end of the semester as well as concept tests developed by the ModularChem Consortium (MC^2), a NSF-funded systemic change project working to change undergraduate chemistry education. The concept test included an experimental design item that was used as a measure of student inquiry skills. In addition, groups of students working in each lab environment were videotaped throughout the semester.

Survey

First, in order to determine student perceptions of their laboratory environment, an end-of-semester survey was administered to a subset of MORE and standard group students. On this

survey, students were given the list of learning goals such as thinking about the meaning of data, performing laboratory techniques, and learning how chemistry research is done. They were asked to respond on a Likert scale to the question: "How do you think Professor X would rank the importance of each of the following goals for learning in the laboratory?" Next, students were asked to rank the same items in terms of what they thought they had learned in the laboratory.

Videotape

Second, groups of students working together on a given laboratory experiment were videotaped and audiotaped each week. Videotapes following the progress of one group in each laboratory setting were transcribed to study the interactions among the students and between the students and the graduate student instructor (GSI). Week 6 of the Control curriculum was deemed as a typical standard lab experience (Control Procedures). By this point in the semester, one would expect the students to have established a pattern of interacting with each other and with their instructor. As a comparison, week 4 of the MORE curriculum (corresponding to the first week of the Stomach Module) was selected because the nature of the experiment was most similar to the Control experience; that is, students were given an experimental question and procedures to carry out (Mod2 Procedures). Week 6 of the MORE curriculum (Mod2 Design), corresponding to the experimental design project in the Stomach Module, was analyzed to document the types of interactions that were promoted by the increased inquiry responsibility.

Videotapes of the lab sessions were transcribed, and the transcripts of the students and instructors speaking were then evaluated to generate a coding scheme that captured the classroom interactions between the lab instructor and the students. Specifically, the instructor-student interactions were parsed into three categories: low-level procedures, mid-level techniques, and high-level reflection. An exchange between the lab instructor and a student(s) was defined as one response or a series of responses concerning one topic area.

Level of interaction	Examples
Low level procedures	• location of equipment • clarification of procedures outlined in lab manual
Mid-level techniques	• tutorials for laboratory technique (e.g. pipette, buret)
High level reflection and explanation	• prompt student reflection of the current model • link the applicable concepts to experimental data

Table 2. Levels of GSI/student interaction

Low-level procedures were generally short exchanges where the lab instructor clarified procedures outlined in the lab manual. Mid-level techniques encompassed exchanges where the lab instructor provided assistance in learning new laboratory techniques such as pipetting or using a buret. High-level reflection included interactions where the lab instructor prompted students to reflect on their model of the system or to provide conceptual explanations. The low and mid-level exchanges are a subset of procedural instructor support, while the high level exchanges reflect the cognitive support offered by the instructor. Table 2 presents examples for each level of interaction. The questions that the lab instructor posed to the students were evaluated using a similar categorization scheme.

Experimental design

Finally, in order to assess student inquiry skills, all students in the participating sections were asked to complete an experimental design task as part of a concept test developed by the NSF systemic change project, MC^2. The pre- and post-experimental design tasks

(ED1 and ED2) measured a student's ability to design a protocol to answer a posed experimental problem. The two tasks were both analytical in nature, asking students to propose a procedure to distinguish the amount of material present. For ED1, students were asked to determine the amount of iron present in two competing cereals and investigate the claim that one cereal contained twice as much iron as the other. For ED2, students were asked to determine how much of a spoonful of sugar would dissolve in a salt-saturated solution (all, some, or none of the sugar). For each task, students were given a list of available equipment, including a balance and spectrophotometer, along with some explanation of what the equipment could be used for. More than one correct method was possible to answer the posed problem for each task.

Student's written responses were entered into a database and evaluated anonymously. Students earned scores in each of the following categories: 1) *Protocol development*, which refers to the extent that implementation of the proposed experimental procedure would provide data to answer the posed question; and 2) *Data analysis and explanation*, which reflects on the students' description of the planned analysis and clarification of its potential implications. These coding dimensions are a subset of the inquiry skills inventory devised by Tamir, Nussinovitz, and Friedler (1982). Their Practical Tests Assessment presents twenty-one categories as an inventory of scientific inquiry skills and describes how it may be used to assess practical lab exams. Some of their classifications include formulation of problems/hypotheses, identification of (in)dependent variables, completeness of design, fitness of the experiment for the tested problem, interpretation of observed data, critical examination of results, explanation of findings, and suggestions of ways to continue the investigation. Since ED1 and ED2 were coded and scored similarly, the following description of analyzing ED2 is akin to the analysis of ED1.

Students offered one or a combination of the following methods: gravimetric, visual, and/or spectroscopic. A list of coded items was constructed to capture the students' written responses. For each coded item, a score (0-2) was assigned to each possible code based on the correctness of the response. A score of "2" rewards a complete and valid response; a score of "1" implies that the response contained minor flaws; and a score of "0" indicates a fundamentally flawed answer. Because the two aspects were not necessarily independent of each other, the data analyses and explanations scores were summed to yield the "data analysis and explanation" score (out of 4).

For each possible method, one common item was the student's explanation of the proposed method to distinguish among the three possible outcomes, that is, all, some, or none of the sugar dissolves. Since students were instructed to *Explain how the results of your experiment would answer the question*, they earned points in accordance to accounting for how the data would distinguish among the three outcomes. The *Analysis/Explanation* subscore refers to the extent to which the proposed experimental design would distinguish among the three possible outcomes. If all three outcomes were distinguishable from the proposed implementation, the student was awarded two points. If one or two outcomes were distinguishable, the student earned one point. An invalid or inappropriate design earned zero points.

Figure 1 presents one student's proposed experimental design. Based on the coding scheme, the student was awarded two points for the *Protocol* dimension because he knows the amounts of salt and sugar added to the solution. By filtering the solution, he has separated any sugar that did not dissolve. Thus, subtracting the known mass of NaCl from the total mass of solids obtained after drying will yield the amount of sugar dissolved in the saturated salt solution. This student was awarded three additional points for the *Analysis/Explanation* dimension. To have received the full four points, this student should

have added a sentence stating that the question could be answered by comparing the initial mass of sugar added to the mass of sugar dissolved. Although this response earned a high score based on the coding scheme, it should be noted that the details of the given procedure are very sparse as compared to instructions in a typical laboratory manual.

> Obtain 100mL of distilled water. Dissolve NaCl until water is saturated. Make sure weight (g) of NaCl is known. Weigh a spoonful of sugar and add to solution. Filter solution into weighed beaker to dry over bunsen burner to 5mL. Let dry overnight to weigh beaker. This will give you total dissolved solids. Subtract the weight of NaCl from it to see how much sugar dissolved. [E7]

Figure 1. Proposed experimental design: High-end response

The proposed experimental design presented in Figure 2 represents a lower end gravimetric response. First of all, the student does not start with a saturated salt solution with 1 gram of salt in 100 mL of water. Second, the design fails to dry the filtered solids causing the measured mass to err on the positive side because of the additional water weight. Finally, the student does not stay on task to answer the posed question and introduces an unnecessary variable (temperature) into the experiment. This response earned 0 points for the protocol, and 1 point for the data analysis/explanation for an overall score of 1. (Although the student mentioned using the spectrophotometer, he earned no additional points for this approach based on the coding scheme.)

> 1) Obtain 1 g of sugar 1 g of salt. Place salt 100 mL of distilled water. Use the balance to weigh it.
> 2) Determine the concentration using the spectrometer.
>
> 3) Put the sugar in. Check temperature. Do it once at 25˚C, 30˚C, 50˚C, 100˚C. 4 trials. Use the bunsen burner.
> 4) Filter the water out. Mass the solid remaining. [I8]

Figure 2. Proposed experimental design: Low-end response

Although students could propose any one of three (or a combination of) methods, it was not necessary to collate responses as a function of proposed method. Because the students' responses were assigned scores based on their ability to answer the posed question, the chosen method or combination of methods was not as important as being able to propose a protocol that would distinguish among the three possible outcomes (i.e., all, some, or none of the sugar dissolves).

Results

Survey

One component of the instructional design was the exploration of concepts through authentic scientific inquiry. Thus, the students were asked about what they thought they learned from the laboratory. To establish a baseline, students were first asked about what they thought the lecture professor wanted them to learn from the laboratory. The students were given nine goals and asked to rate them on a Likert scale from 1 (strongly disagree) to 5 (strongly agree). (Table 3 lists the goals in order of importance with the overall average.)

There were no significant differences in student response on what the MORE and Control students believed the professor thought were important goals for learning in the laboratory, except in the area of relating chemistry to real-world problems. The Control students rated this item significantly higher than the MORE students ($p=0.042$, two-tailed unpaired t-test).

Students should learn...	Control Mean	MORE Mean	p-value[b]
1...to understand chemistry concepts better through hands-on experience.	4.66[c]	4.38[d]	0.18
2...to think about the meaning of data.	4.37[c]	4.46[d]	0.62
3...to relate chemistry to real-world problems.	4.35[e]	3.85[d]	0.04
4...to present my understanding of chemistry to others (written or oral).	3.63[e]	3.66[f]	0.90
5... to find an answer to an experimental question.	3.51[e]	3.36[d]	0.55
6...to work together in groups.	3.38[e]	3.47[f]	0.71
7...how chemistry research is done.	3.21[e]	3.50[f]	0.22
8...to perform laboratory techniques.	3.29[e]	3.33[d]	0.83
9...to apply mathematical relationships (formulas) to an experimental system.	2.91[e]	2.74[f]	0.50

Table 3. How do you think the professor would rank the importance of the following goals for learning in the laboratory? (Survey 2)
[a]evaluated on a scale from 1 (strongly disagree) to 5 (strongly agree); [b] calculated from two-tailed unpaired t-tests; [c] n=35; [d] n=39; [e] n=34; [f] n=38

However, there were significant differences when comparing the MORE and Control students' beliefs about they learned in the laboratory. First, there is an interesting discrepancy between what students perceive to be important in the professor's eyes as opposed to what they think they actually learned in the laboratory. The MORE students rated the same top five learning goals for what they thought they had learned and what they perceived the professor wanted them to learn. In contrast, the Control students rated laboratory technique and applying formulas 1st and 3rd, respectively, as compared to 6th and 9th, for what they perceived the Professor wanted them to learn.

Secondly, the MORE students rated five of the nine goals significantly higher in terms of what they thought they had learned in the laboratory. (See Table 4.) The only aspects the Control students rated more highly were learning to perform laboratory techniques and applying mathematical relationships. No differences were measured for working together in groups and finding an answer to an experimental question. Here, no differences were expected since students worked in groups in both the MORE and Control laboratory and since experiments for both sought to answer posed experimental questions.

Third, the MORE students strongly rated four of the nine items (average score greater than 4). These four included understanding chemistry concepts, thinking about the meaning of data, relating chemistry to the real world, and presenting one's understanding of chemistry to others. In contrast, the Control students only ranked one item with an average score

greater than 4, specifically, learning to perform laboratory techniques. These results illustrate the different perceptions of what the MORE and Control students felt they learned from their laboratory experience.

I learned...	Control Mean	MORE Mean	p-value[b]	Effect Size[c]
1...to understand chemistry concepts better through hands-on experience.	3.67[d]	4.26[e]	<0.01	57%
2...to think about the meaning of data.	3.92[d]	4.54[e]	<0.01	66%
3...to relate chemistry to real-world problems.	3.77[g]	4.33[e]	0.024	50%
4...to present my understanding of chemistry to others (written or oral).	2.72[d]	4.24[h]	<0.001	160%
5... to find an answer to an experimental question.	3.67[d]	3.95[e]	0.18	32%
6...to work together in groups.	3.64[d]	3.92[h]	0.27	24%
7...how chemistry research is done.	2.97[d]	3.59[e]	0.01	58%
8...to perform laboratory techniques.	4.28[d]	3.92[h]	0.070	-46%
9...to apply mathematical relationships (formulas) to an experimental system.	3.78[d]	2.95[h]	<0.001	-82%

Table 4. Student perception of what they learned in the laboratory
[a] evaluated on a scale from 1 (strongly disagree) to 5 (strongly agree), [b] calculated from two-tailed unpaired t-tests; [c] calculated by dividing the mean difference by the standard deviation for the Control group; [d] n=36; [e] n=39; [g] n= 35; [h] n=38

The list of learning goals in Table 4 contains three inquiry-related measures: learning to think about the meaning of data, learning to present one's understanding of chemistry, and learning how chemistry research is done. Each of these aspects was rated significantly higher by the MORE students than the Control. Thus, the survey data indicates that the MORE students believe their laboratory environment supported the inquiry process and fostered their development of inquiry skills.

In summary, the survey results indicate that although the MORE and Control students shared similar perceptions regarding the professor's goals for the laboratory, their actual laboratory experiences were distinct. Specifically, the MORE students felt that their lab experience focused on understanding chemistry concepts, thinking about the meaning of data, and relating chemistry to real-world problems. In comparison, the Control students felt they learned to perform laboratory techniques and to apply algorithms to an experimental system. These results indicate that the MORE curriculum and instruction provided students with a more authentic scientific inquiry experience.

Videotape

Videotapes were analyzed with respect to lab instructor/student group interactions and within group student/student interactions to determine the extent to which students were reflective during the laboratory period. Specifically, the following questions were explored in the videotape analyses:

1) How are the classroom interactions (GSI/student and within student/student) different when the students are completing prescribed procedures? This compares the interactions in Control (Wk 6) to MORE (Wk 4). The experiment completed in the first week of the Stomach Module (Mod2 Procedures) is most similar to the typical experiments in the Control laboratory (Control Procedures) as students are following a set of prescribed procedures to determine the acid-neutralizing capacity of an antacid.

2) How do the classroom interactions progress across a module as students engage in a multi-week exploration of a system? This targets the interactions in MORE (Wk 4), Mod2 Procedures, as compared to MORE (Wk 6), Mod2 Design. The same group of students videotaped completing the prescribed procedures were taped two weeks later implementing an experiment they had designed to further their understanding of the chemistry of the stomach.

In order to gain a clearer picture of the differences between the MORE and Control sessions, a frequency distribution was constructed to represent the occurrence of each level of interaction across 10-minute intervals for each laboratory session. That is, the number of low level exchanges were summed for each 10-minute period, the number of mid-level exchanges, and so on, for the first two hours of instruction. These frequency counts of the different levels were entered and counts for different levels during the same 10-minute period were stacked. In addition to the GSI/student interactions, the incidences of within student/student interactions that indicated a high level of reflection among the students were also included in the frequency distribution. For example, when students related their protocol to their model of the system or discussed the implications of their data on their conceptual understanding, these interactions were classified as "student reflection".

1) How are the classroom interactions different when the students are completing prescribed procedures?

Figure 3 presents a graphical representation of the GSI/student interactions to compare classroom interactions when students are provided with procedures to complete. The low and mid-level interactions are evenly spread throughout the two hour period for both the Control Procedures and the Mod2 Procedures sessions. The interesting difference lies in the absence of any high level interactions in the Control class.

First, the MORE students engaged in two high level interactions with the graduate student instructor, or GSI (time 40, time 80). For instance, in the following exchange, the instructor prompted the students to consider the implications of their titration data.

> Student: *It figures. The rate I'm going* (refers to rate of titration). *What would cause us to need more (NaOH)?*
> GSI: *So does this mean you have a high ANC or low ANC? You're adding base right now.* (time 1:29:55)

At this point in the experiment, the students were titrating NaOH (known concentration) into their solution of antacid and excess HCl (known volume and concentration). (The goal of this exercise was for students to determine the acid neutralizing capacity (ANC) of their assigned antacid formulation.) In the exchange, the student asked the GSI why more

NaOH would be needed. The GSI seized this opportunity to have the students engage in sensemaking and consider the implications of NaOH titration on the ANC magnitude of their antacid formulation.

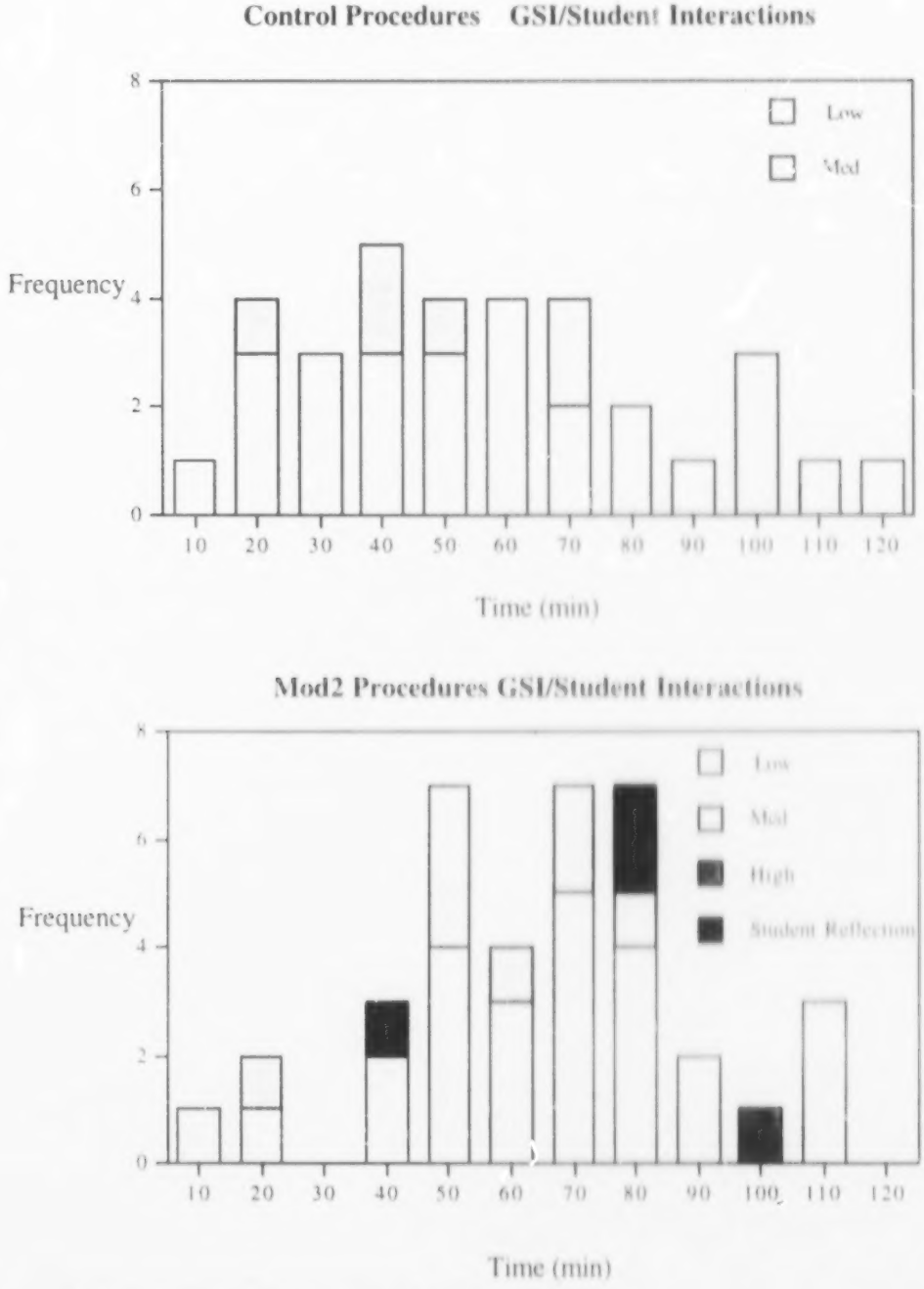


Figure 3. Control Procedures vs. Mod2 Procedures

In addition, the within student/student interactions for the Mod2 Procedures session revealed two additional high level exchanges among the students as they reflected on what they were doing and the potential implications. For instance, the MORE students stopped themselves to ask, "Does it matter...?" and a discussion ensued regarding the setup of the experiment. The presence of these high level exchanges that are reflective in nature suggest that the MORE students have a better understanding of the purpose and/or goals of the experiment, while the Control students focus primarily on following the laid-out recipe with little thought of the underlying concepts or overarching purpose. Furthermore, the student reflection in the MORE classroom may imply that the students have internalized the MORE Cycle and are explicitly practicing the Reflect aspect of the cycle during the lab period.

However, several alternative explanations could account for the differences observed for the Mod2 Procedures and Control Procedures classes. First, one may argue that the differences in high level reflection were curriculum-dependent, that is, the Control lab curriculum did not provide potential opportunities for the instructor to prompt students to reflect unlike the MORE curriculum. However, two of the interactions between the Control instructor and students could easily have been transformed into high level exchanges. For instance, at time 45, the students asks "Is that enough for titration?" The instructor could have capitalized on this moment and responded, "Do you expect to need more or less than the previous titration done at a lower temperature? Why?" Thus, in addition to determining the point at which the students are currently at in the protocol (as the Control GSI did), the instructor can return the responsibility to the students to reflect and make some sense of how solubility changes with temperature. Another potential moment to ask a high level question occurred at time 1:18. Here, the student asks "Are we allowed to add water to our borax solution?" Here, the instructor could have asked the students "Does it matter? Why?" in order to prompt the students to think about the implications of adding water.

A second argument for the observed differences is that the Control students were reflecting quietly to themselves rather than verbalizing their thoughts to their classmates. If this were the case, the analysis would not be able to detect any such personal reflection. However, student response to the interview question, *"During lab while you were doing experiments, did you find yourself trying to understand what was going on in the experiment?"*, provides some anecdotal evidence that Control students were not engaging in reflective practices during the lab period. Two responses articulate the general sentiment of the Control students:

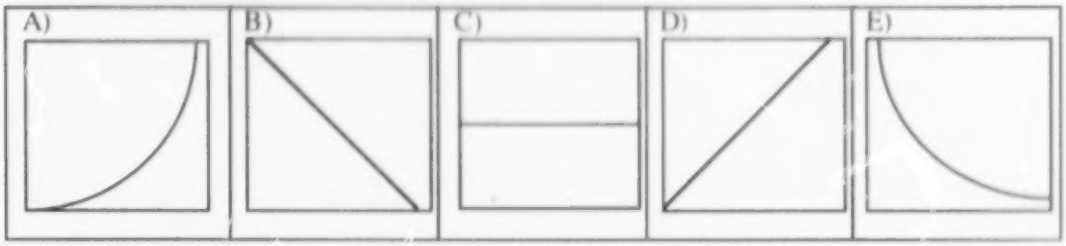
> *Student 1: Unfortunately, I'd probably say no—because the experiment was laid out that wasn't necessary of you. You just had to do the experiment and write down the numbers they told you to write down. So you pretty much were doing the experiment in order to get the results, not necessarily to understand the results.*
>
> *Student 2: We're told how to do the calculations but there was never much discussion about "Well, this is what the connection is to what you are learning."*

These responses indicate that the predefined protocols and analyses typical of the standard laboratory experience were not prompting the Control students to try and understand what they were doing and how it related to the underlying concepts.

A third argument is that the ensuing laboratory report would prompt the Control students to reflect. The primary goal of the Control experiment was to explore the van't Hoff relationship and plot the resulting lnK vs. 1/T line. Supposing the Control students grappled with this concept, one would assume that they would do well on an exam

question assessing their understanding of this thermodynamics concept. The following question, analogous to the Control lab experience, was included on the comprehensive final exam at the end of the course:

...choose which of the following five graphs best describes the behaviors listed below: lnK as a function of 1/T for an endothermic reaction.



The expectation is that the Control students, who had this concept reinforced in the laboratory, should perform as well or better than the MORE students, who had seen this concept in lecture but had not performed any thermodynamics experiment. As Table 5 indicates, 60% of the Control students did select the correct answer (B). However, 77% of the MORE students answered correctly. Although this is not a significant difference ($p=0.14$, two-tailed Fisher's exact test), the fact that the Control lab experience did not translate into a higher percentage of students answering correctly implies that the Control students finished the lab report without a deeper understanding of the underlying lnK vs. 1/T relationship.

	Answer A	Answer B	Answer D	Answer E
Control	5.3%	60.5%	26.3%	7.9%
MORE	2.6%	76.9%	7.7%	12.8%

Table 5. Control vs. MORE Performance on van't Hoff exam question

An additional 26% of the Control students chose the positively-sloped line response (D). This implies that many of the Control students recalled the linear relationship from their lab experience, but again, lacked a conceptual understanding of the relationship. In comparison, the second most popular response by the MORE students was answer (E), indicating that these students possessed a qualitative understanding of the relationship between lnK and 1/T. The MORE students were more likely to understand the qualitative relationship, that is, as temperature increases, lnK also increases ($p=0.026$, two-tailed Fisher's exact test).

In summary, the observed differences between the Control Procedures and Mod2 Procedures sessions seem to indicate that (a) the MORE students internalized the Reflect aspect of the MORE Cycle and engaged in reflective practice during the lab period; and (b) the MORE instructor prompted the students to make sense of what they were doing.

2) How do the classroom interactions progress across a module as students engage in a multi-week exploration of a system?

The second question explores the implications of sustained research investigation on the classroom interactions. Figure 4 illustrates the dramatic differences between the two class sessions, contrasting the same group of students in the first and third weeks of Module 2. First, the low level interactions are considerably reduced during the students' independent project, while the mid-level interactions are essentially eliminated. Second, the number of high level exchanges nearly triples from two in Week 4 to seven in Week 6. Finally, the seven instances of student reflection indicate that the students themselves are monitoring

their own understanding of their project as they relate their protocol to the experimental purpose or current model of the system. (See Table 6 for examples of student reflection.)

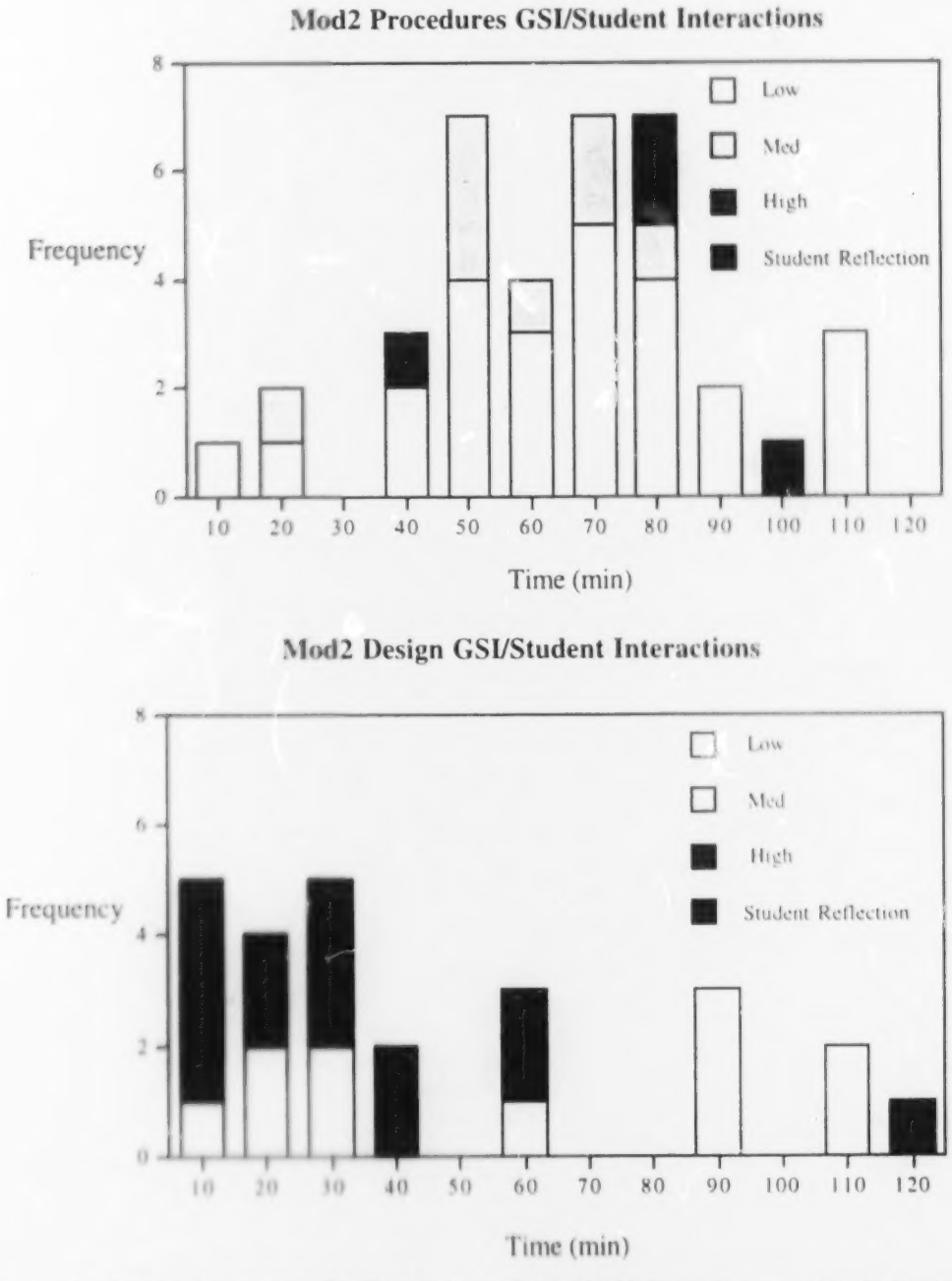


Figure 4. Progression across the Stomach Module: Mod2 Procedures vs. Mod2 Design

It is interesting to note that the majority of these high level exchanges for the GSI/student and student/student interactions occur primarily in the first hour of the lab period. One possible explanation for this observation is that students invest considerable time at the

beginning of the laboratory making sure they understand the reasons behind what they are doing and why they are doing it. These graphs offer compelling evidence that students are much more reflective when they carry out procedures that they themselves have constructed. Not only do they have to make sense of why they do a certain step, they must also consider the relevance of the resulting data to their posed experimental question. In comparison, when students follow a given set of procedures, they focus their attention on completing the steps rather than understanding the rationale behind them.

24	f3: *I'm trying to think—why did we add*—While you're doing that, I'm thinking this out—You know what I think? So instead of adding 100mL, maybe we should add 50. Should we add 50? Like you said to simulate parietal cells (cells that secrete acid).
28	f3: *Did we do this because—I'm trying to figure out why we added—? Why are we choosing—I'm trying to figure out*—Do you remember why we chose 0.16 grams for 50 mL?
37	m3: Why don't I measure you—We're going to start with 100-mL. f3: How many minutes are we going? 90 minutes? 90 times 2— m3: Yeah— f3: —that's 180 ml. m3: Yeah, that's way too much. f3: Well, the stomach has 100mL fasting in it so 280mL isn't too much.
1:06	f3: *Why are we keeping track of pH*? m3: Just as another measure. Just to see because remember the second experiment where the pH...
2:13	f3: You know, it's just occurring to me—*We're not sure how much relief this is giving—like how much this is digesting*—
2:35	m3: *Is the time element big in the filtration*? Like do you need to immediately do it? f3: I think it's just for breakdown time like if you're gonna say you did it for an hour and you're gonna compare the food then you wanna do it as soon as to an hour as you can. Do you know what I'm saying?

Table 6 MORE student reflections (italics added for emphasis)

In summary, the sustained experimentation allows students to learn relevant techniques and become familiar with the experimental system enabling independent study of the system in the later weeks of each module.

Summary of videotape analyses

When viewing the student interactions as snapshots of each type of class session, the following list emerges in order of which treatment promoted the most high level interactions and student reflection: Mod2 Design > Mod2 Procedures > Control Procedures. These results make sense in light of the instructional framework; the three aspects of the framework were (a) exploration of concepts through authentic inquiry; (b) promotion of metacognition; and (c) support for guided discovery. (See Table 7.)

On the low end, the Control Procedures experience did little to address any of the three aspects of the framework. Next, Mod2 Procedures, which incorporated the MORE Cycle and some level of scaffolding (addressing promotion of metacognition and support for guided discovery, respectively), promoted student reflections and also fostered a number of high level interactions between the students and the instructor. Finally, the Mod2 Design integrated metacognition (through the use of the MORE Cycle and instructor prompts), authentic inquiry (through the sustained investigations and design experiences), and support for guided discovery (through the increased inquiry responsibility from one module to the next). This experience yielded the greatest number of student reflections and

high level interactions between the students and the instructor. Consequently, these snapshots of students and their instructor speaking during the laboratory indicate the value of providing an authentic inquiry experience, promoting metacognition, and supporting guided discovery.

Laboratory	Framework Aspects Addressed	Instructional implementation
1) Mod2 Design	a) Exploration through authentic inquiry b) Promotion of metacognition c) Support for guided discovery	a) Experimental design b) MORE Cycle c) Increased inquiry responsibility across modules
2) Mod2 Procedures	a) Promotion of metacognition b) Support for guided discovery	a) MORE Cycle b) Instructor scaffolding
3) Control Procedures	None	N/A

Table 7. Summary of videotape analyses

Experimental design

Statistical analyses were carried out for all written responses for the two experimental design assessments (ED1 and ED2). The maximum possible score for both ED1 and ED2 was 6. Table 8 shows that the MORE and Control students performed equivalently on ED1. This indicates that the two groups of students did not begin the semester with measurably different experimental design skills. However, the MORE students did significantly better ($p=0.037$, two-tailed unpaired t-test) than the Control students on ED2 with a mean difference of 0.91. The MORE students scored significantly higher only on the Analysis/Explanation dimension ($p=0.026$, two-tailed unpaired t-test). (Note: Because fewer Control students completed ED1, including ED1 scores as a covariate would result in a loss of many degrees of freedom. Thus, ED1 scores were not used in ANCOVA analyses or as a predictor for regression analyses.)

	Control Mean	MORE Mean	p-value	Effect Size
Experimental Design 1 Score (6)	1.82[a]	2.00[b]	0.57	N/A
Protocol (2)	0.96	1.14	0.61	N/A
Analysis/Explanation (4)	0.86	0.86	0.99	N/A
Experimental Design 2 Score (6)	1.65[c]	2.56[d]	0.037	56%
Protocol (2)	0.46	0.67	0.23	32%
Analysis/Explanation (4)	1.19	1.90	0.026	59%

Table 8. Student performance on experimental design assessments
[a] n=22; [b] n=36; [c] n=37; [d] n=39

Regression analyses show that Lab (More vs. Control) and Math SAT combine to contribute significantly to the model for ED2 scores (F-ratio=3.96; p=0.023). While gender and previous chemistry background did not serve to predict ED2 performance, Lab and Math SAT accounted for 7% of the variance seen in ED2 scores (adjusted R^2=0.073). (See Table 9 for the summary of regression analyses.) With this regression model, the effect size associated with the Lab is still high at 50% (calculated by dividing the coefficient by the standard error of estimate). The individual contribution of Lab to ED2 performance indicates that the MORE group scores significantly higher than the Control (p=0.033), while the individual contribution of Math SAT is marginal at predicting ED2 performance (p=0.074).

However, it should be noted that despite these significant differences between groups, only one in four of the MORE students (9 of 39) gave complete and correct experimental protocols at the end of the semester as compared to one in nine of the Control students (4 of 38).

Variable	Coefficient	Std Coefficient	t-value	p-value
Intercept	-2.717	0.000	-1.121	0.266
Lab	0.924	0.241	2.172	0.033
Math SAT	0.006	0.202	1.816	0.074
R: 0.313; R^2: 0.098; Adjusted R^2: 0.073; Std error of estimate: 1.853				

Table 9. Regression summary results of ED2 performance

These results suggest that the increased inquiry responsibility and emphasis on explanations in the MORE curriculum help students learn to design appropriate experimental protocols and data analyses, and present coherent explanations. Development of superior experimental design skills in the MORE students compared with Control students is also consistent with student perceptions of their laboratory experiences (as seen in the survey responses.)

Discussion

Analyses of survey responses, student-instructor discourse, and experimental design measures showed some interesting differences. For the survey items, both groups perceived that the course professor would rank goals related to conceptual understanding and making sense of data highly and rank performing laboratory techniques and applying mathematical relationships lowly. However, in terms of what the students believed they learned in the laboratory, the MORE students rated five of the nine items significantly higher, including items relating to understanding and presenting chemistry ideas, and inquiry-related items such as thinking about the meaning of data and learning how chemistry research is done.

Analysis of the videotapes showed a marked contrast between the two laboratory environments. First of all, the interaction in the MORE environment focused on a reflective level as the instructor prompted students to consider their model or the overarching goal of the experiment, while the standard instructor emphasized procedural details and proper laboratory technique. Second, while standard student conversations demonstrate their lack of a coherent model of what they are doing and why they are doing it, the MORE students frequently connect their actions with their current model. These results are not surprising; these differences are consistent with the distinct goals and philosophy of each laboratory environment. That is, these results show that the MORE curriculum and instruction is successful at promoting metacognition and engaging students in more authentic inquiry.

The different classroom experiences promoted significantly different outcomes on inquiry assessments. The experimental design results suggest that the increased inquiry

responsibility and emphasis on explanations in the MORE curriculum help students learn to design appropriate experimental protocols and data analyses, and present coherent explanations. Development of superior inquiry skills in the MORE students compared with standard students is also consistent with student perceptions of their laboratory experiences as seen in the survey responses. These findings imply that the laboratory curriculum and instruction can be structured such that significant gains can be achieved in student inquiry skills.

Several perspectives may contribute to understanding the occurrence of the high level of reflection and modeling in the experimental design weeks. First, from the information processing perspective, the higher cognitive load placed on a student in learning new laboratory techniques and exploring an unfamiliar system potentially overwhelms the student, and thus prevents the student from engaging in such high level thought. Second, from the cognitive apprenticeship model, students have been scaffolded to use the MORE Cycle as a tool to guide their thinking in the laboratory. Because students are engaged in authentic inquiry; they will encounter instances during their explorations that prompt them to engage in high level reflection similar to the experiences of research scientists. For instance, students responsible for designing their experiments must determine what data is appropriate for answering the experimental question and how that data should be analyzed. As students adopt the MORE Cycle, they adopt the Reflect aspect of the Cycle during the laboratory and are prompted to Explain following their practical experience.

The findings from the inquiry assessments are both encouraging and somewhat disturbing. While the authentic inquiry experience seemed to foster the MORE students' design skills, it is disappointing to see that all of the students completing one semester of college chemistry do so poorly on the experimental design assessments. The experimental design task was not a complex problem that required a deep understanding of chemistry. It did, however, require students to consider how to answer the posed problem. While the MORE experience causes some positive effect, only 25% of the MORE students designed a completely acceptable protocol (as compared to 10% of the Control students). If one of the instructor's goals for the classroom is fostering experimental design skills, the students not only need opportunities to design their own experiments throughout the semester, but also need a great deal of feedback and support during this process. In addition, I believe it is important to make explicit to students the criteria on which their designs will be evaluated so that students can become more critical of their own designs and those of their peers. This reflective, self-assessment process has been successful in pre-college science classes (White and Frederiksen, 1995).

I believe the evidence gathered in the MORE project bear witness to the potential of educational research to make the school experience more meaningful and productive. Embedding the MORE Cycle into the curriculum and instruction gave students guidelines for how to think during the laboratory, that is, applying the cycle compelled students to reflect and explain what they were doing and why they were doing it. The videotape analyses provided snapshots of the MORE students explicitly integrating the Reflect portion of the cycle while they were carrying out their lab experiments. The MORE Cycle in conjunction with experimental design and scaffolding yielded the greatest number of cognitive exchanges among the students as well as between the instructor and the students. Finally, the authentic inquiry experience appears to have fostered the development of inquiry skills, specifically, experimental design skills. By situating students in an environment that more closely resembles expert practice, instruction can promote scientific inquiry.

References

Arons, A.B. (1993). Guiding insight and inquiry in the introductory physics laboratory. *Physics Teacher*, 31, 5, 278-282.

Baird, J.R., Fensham, P.J., Gunstone, R.F., and White, R.T. (1991). The importance of reflection in improving science teaching and learning. *Journal of Research in Science Teaching*, 28, 2, 163-182.

Brown, A.L., Bransford, J.D., Ferrara, R.A., and Campione, J.C. (1983). Learning, remembering, and understanding. In P.H. Mussen, (Ed.), *Handbook of Child Psychology*, vol. 3: *Cognitive Development.*. New York: Wiley.

Brown, A.L. & Campione, J.C. (1993). Guided discovery in a community of learners. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice*. Cambridge, MA: MIT Press/Bradford Books.

Carey, S., Evans, R., Honda, M., Jay, E., and Unger, C. (1989). 'An experiment is when you try it and see if it works': A study of grade 7 students' understanding of the construction of scientific knowledge. *International Journal of Science Education*, 11, 514-529.

Collins, A., Brown, J.S., and Holum, A. (1991). Cognitive apprenticeship: Making thinking visible. *American Educator* (Winter), 6-11, 38-46.

Collins, A., Brown, J.S., and Newman, S.E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics, In L.B. Resnick (Ed.), *Knowing, Learning, and Instruction.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Flavell, J.H. (1976). Metacognitive aspects of problem solving. In L.B. Resnick's (Ed.) *In the Nature of Intelligence*. Lawrence Erlbaum Associates: Hillsdale, NJ.

Funk, H.J., Okey, J.R., Fiel, R.L., Jans, H.H., and Sprague, C.S. (1979). *Learning Science Process Skills*. Duboque, IA: Kendall/Hunt.

Gottfried, S., Hoots, R., Creek, R., Tamppari, R., Lord, T., and Sines, R. (1993). College biology teaching: A literature review, recommendations, and a research agenda. *The American Biology Teacher,* 55, 6, 340-348.

Koslowski, B. (1996). *Theory and Evidence: The Development of Scientific Reasoning*. Cambridge, MA: The MIT Press.

Kreitler, H. and Kreitler, S. (1974). The role of the experiment in science education. *Instructional Science*, *3*, 75-88.

Kuhn, D. (1993). Science as argument: Implications for teaching and learning scientific thinking. *Science Education*, 77(3), 319-337.

Lloyd, B.W. and Spencer, J.N. (1994). New directions for general chemistry. *Journal Of Chemical Education* , 71, 3, 206-209.

Padilla, M.J. (1991). Science activities, process skills, and thinking. In S.M. Glynn, R.H. Yeany, and B.K. Britton (Eds.), *The Psychology of Learning Science*. Lawrence Erlbaum Associates: Hillsdale, NJ.

Schauble, L., Glaser, R., Duschl, R.A., Schulze, S., and John, J. (1995). Students' understanding of the objectives and procedures of experimentation in the science classroom. *Journal of the Learning Sciences*, 4, 2, 131-166.

Spencer, J.N. (1993-1994). The general chemistry curriculum: Different times, different students, same course. *Journal of College Science Teaching*, 23, 3, 159-161.

Strike, K.A. and Posner, G.J. (1992). A revisionist theory of conceptual change. In R.A. Duschl & R.J. Hamilton, (Eds.), *Philosophy of Science, Cognitive Psychology, and Educational Theory and Practice*. Albany: State University of New York Press.

Tien, L.T., Rickey, D., and Stacy, A.M. (in press). The MORE Cycle: A structure to guide students' thinking in the laboratory. *Journal of College Science Teaching*.

Tien, L.T. and Stacy, A.M. (1998). Cognitive apprenticeship in the chemistry classroom. Presented at the American Chemical Society Conference. Dallas, TX: ACS.

Tobin, K. (1990). Research on laboratory activities: In pursuit of better questions and answers to improve learning. *School Science and Mathematics*, 90, 403-418.

Welch, W.W. (1979). Twenty years of science curriculum development: A look back. In D.C. Berliner (Ed.), *Review of Research in Education*., Ch. 7. American Educational Research Association.

White, B.Y. & Frederiksen, J.R. (1995). *The ThinkerTools Inquiry Project: Making Scientific Inquiry Accessible to Students and Teachers*. CM-95-02. U.C. Berkeley, School of Education, Causal Models Research Group.

Yap, K.C. and Yeany, R.H. (1988). Validation of hierarchical relationships among Piagetian Cognitive Modes and integrated science process skills for different cognitive reasoning levels. *Journal of Research in Science Teaching*, 25, 247-281.